

STRUCTURED PSEUDOSPECTRA AND THE CONDITION OF A  
NONDEROGATORY EIGENVALUE\*

MICHAEL KAROW†

**Abstract.** Let  $\lambda$  be a nonderogatory eigenvalue of  $A \in \mathbb{C}^{n \times n}$  of algebraic multiplicity  $m$ . The sensitivity of  $\lambda$  with respect to matrix perturbations of the form  $A \rightsquigarrow A + \Delta$ ,  $\Delta \in \Delta$ , is measured by the structured condition number  $\kappa_\Delta(A, \lambda)$ . Here  $\Delta$  denotes the set of admissible perturbations. However, if  $\Delta$  is not a vector space over  $\mathbb{C}$ , then  $\kappa_\Delta(A, \lambda)$  provides only incomplete information about the mobility of  $\lambda$  under small perturbations from  $\Delta$ . The full information is then given by the set  $K_\Delta(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\| \leq 1\} \subset \mathbb{C}$  that depends on  $\Delta$ , a pair of normalized right and left eigenvectors  $x, y$ , and the norm  $\|\cdot\|$  that measures the size of the perturbations. We always have  $\kappa_\Delta(A, \lambda) = \max\{|z|^{1/m}; z \in K_\Delta(x, y)\}$ . Furthermore,  $K_\Delta(x, y)$  determines the shape and growth of the  $\Delta$ -structured pseudospectrum in a neighborhood of  $\lambda$ . In this paper we study the sets  $K_\Delta(x, y)$  and obtain methods for computing them. In doing so we obtain explicit formulae for structured eigenvalue condition numbers with respect to many important perturbation classes.

**Key words.** eigenvalues, structured perturbations, pseudospectra, condition numbers

**AMS subject classifications.** 15A18, 15A57, 65F15, 65F35

**DOI.** 10.1137/070695836

**Notation.** The symbols  $\mathbb{R}$  and  $\mathbb{C}$  denote the sets of real and complex numbers, respectively.  $\mathbb{K}^{m \times n}$  is the set of  $m \times n$  matrices and  $\mathbb{K}^n = \mathbb{K}^{n \times 1}$  is the set of column vectors of length  $n$ ,  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ . By  $A^\top$ ,  $\bar{A}$ ,  $A^*$ ,  $\Re A$ ,  $\Im A$ , and  $\text{tr}(A)$  we denote the transpose, the conjugate, the conjugate transpose, the real and the imaginary parts, and the trace of  $A \in \mathbb{C}^{m \times n}$ . Furthermore,  $I_n$  stands for the  $n \times n$  unit matrix. Finally,  $\underline{n} = \{1, \dots, n\}$  for any positive integer  $n$ .

**1. Introduction.** The subject of this paper are the sets

$$(1.1) \quad K_\Delta(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\| \leq 1\}, \quad x, y \in \mathbb{C}^{n \times n},$$

where  $\|\cdot\|$  is a norm on  $\mathbb{C}^{n \times n}$  and  $\Delta \subseteq \mathbb{C}^{n \times n}$  is assumed to be a closed cone (the latter means that  $\Delta \in \Delta$  implies  $r\Delta \in \Delta$  for all  $r \geq 0$ ). Our motivation for considering these sets stems from eigenvalue perturbation analysis by means of pseudospectra. The sets  $K_\Delta(x, y)$  provide the full first order information about the sensitivity of a nonderogatory eigenvalue with respect to structured matrix perturbations. This is explained in some detail in the following discussion.

Let  $\lambda \in \mathbb{C}$  be a nonderogatory eigenvalue of algebraic multiplicity  $m$  of  $A \in \mathbb{C}^{n \times n}$ . Let  $x \in \mathbb{C}^n \setminus \{0\}$  be a right eigenvector, i.e.,  $Ax = \lambda x$ . Then there exists a unique left generalized eigenvector  $\hat{y} \in \mathbb{C}^n \setminus 0$  satisfying

$$\hat{y}^*(A - \lambda I_n)^m = 0, \quad \hat{y}^*(A - \lambda I_n)^{m-1} \neq 0, \quad \hat{y}^*x = 1.$$

Let  $y^* = \hat{y}^*(A - \lambda I_n)^{m-1}$ , and let  $\|\cdot\|$  be an arbitrary norm on  $\mathbb{C}^{n \times n}$ . Under a small perturbation of  $A$  of the form

$$(1.2) \quad A \rightsquigarrow A(\Delta) = A + \Delta, \quad \Delta \in \mathbb{C}^{n \times n},$$

\*Received by the editors June 29, 2007; accepted for publication (in revised form) by M. Embree August 3, 2010; published electronically November 30, 2010.

<http://www.siam.org/journals/simax/31-5/69583.html>

†Matheon, Technische Universität Berlin, D-10623 Berlin, Germany (karow@math.TU-Berlin.de).

the eigenvalue  $\lambda$  splits into  $m$  eigenvalues  $\lambda_1(\Delta), \dots, \lambda_m(\Delta)$  of  $A(\Delta)$  with the first order expansion [16]

$$(1.3) \quad \lambda_j(\Delta) = \lambda + \theta_j(\Delta) + \mathcal{O}(\|\Delta\|^{2/m}), \quad j \in \underline{m},$$

where  $\theta_1(\Delta), \dots, \theta_m(\Delta)$  are the  $m$ th roots of  $y^* \Delta x \in \mathbb{C}$ . Obviously,

$$|\theta_j(\Delta)| = |y^* \Delta x|^{1/m} = \mathcal{O}(\|\Delta\|^{1/m}), \quad j \in \underline{m}.$$

We assume now that the perturbations  $\Delta$  are elements of a nonempty closed cone  $\Delta \subseteq \mathbb{C}^{n \times n}$ . Let

$$\kappa_{\Delta}(A, \lambda) = \max\{ |y^* \Delta x|^{1/m}; \Delta \in \Delta, \|\Delta\| \leq 1 \}.$$

Then  $\kappa_{\Delta}(A, \lambda)$  is the smallest number  $\kappa$  such that

$$|\lambda_j(\Delta) - \lambda| \leq \kappa \|\Delta\|^{1/m} + \mathcal{O}(\|\Delta\|^{2/m}) \quad \text{for } \Delta \in \Delta.$$

The quantity  $\kappa_{\Delta}(A, \lambda)$  is called the structured condition number of  $\lambda$  with respect to  $\Delta$  and the norm  $\|\cdot\|$ . It measures the sensitivity of the eigenvalue  $\lambda$  if the matrix  $A$  is subjected to perturbations from the class  $\Delta$ . In recent years some work has been done in order to obtain estimates or computable formulae for  $\kappa_{\Delta}(A, \lambda)$  [3, 4, 5, 7, 13, 15, 16, 18, 17, 20, 21, 23]. However, the condition number cannot reveal how the eigenvalue moves in a specific direction under structured perturbations. For instance if  $\lambda$  is a simple real eigenvalue of a real matrix  $A$  and the perturbations  $\Delta$  are also assumed to be real, then the perturbed eigenvalue  $\lambda(\Delta)$  remains on the real axis if  $\|\Delta\|$  is small enough. Information of this kind can be obtained from the structured pseudospectrum  $\sigma_{\Delta}(A, \epsilon)$ , which is defined as follows [6, 12, 24]:

$$\sigma_{\Delta}(A, \epsilon) = \{z \in \mathbb{C}; z \text{ is an eigenvalue of } A + \Delta \text{ for some } \Delta \in \Delta, \|\Delta\| \leq \epsilon\}, \quad \epsilon > 0.$$

Let  $\mathcal{C}_{\Delta}(A, \lambda, \epsilon)$  denote the connected component of  $\sigma_{\Delta}(A, \epsilon)$  that contains the eigenvalue  $\lambda$ . Then we have for sufficiently small  $\epsilon$  that

$$\mathcal{C}_{\Delta}(A, \lambda, \epsilon) = \{\lambda_j(\Delta); \Delta \in \Delta, \|\Delta\| \leq \epsilon, j \in \underline{m}\}.$$

We now consider the sets

$$(1.4) \quad K_{\Delta}^{(m)}(x, y) = \{z \in \mathbb{C}; z^m \in K_{\Delta}(x, y)\}.$$

In other words,  $K_{\Delta}^{(m)}(x, y)$  is the set of all  $m$ th roots of the numbers  $y^* \Delta x$ , where  $\Delta \in \Delta$ ,  $\|\Delta\| \leq 1$ . We have

$$(1.5) \quad \kappa_{\Delta}(A, \lambda) = \max\{|z|; z \in K_{\Delta}^{(m)}(x, y)\}$$

$$(1.6) \quad = \max\{|z|^{1/m}; z \in K_{\Delta}(x, y)\}.$$

Moreover, (1.3) yields that

$$(1.7) \quad \lim_{\epsilon \rightarrow 0} \frac{\mathcal{C}_{\Delta}(A, \lambda, \epsilon) - \lambda}{\epsilon^{1/m}} = K_{\Delta}^{(m)}(x, y),$$

where the limit is taken with respect to the Hausdorff metric. More explicitly, (1.7) states that for each  $\delta > 0$  there exists an  $\epsilon_0 > 0$  such that, for all positive  $\epsilon < \epsilon_0$ ,

$$(1) \quad \mathcal{C}_{\Delta}(A, \lambda, \epsilon) \subset \lambda + \epsilon^{1/m} \mathcal{U}_{\delta}(K_{\Delta}^{(m)}(x, y)),$$

$$(2) \quad \lambda + \epsilon^{1/m} K_{\Delta}^{(m)}(x, y) \subset \mathcal{U}_{\delta}(\mathcal{C}_{\Delta}(A, \lambda, \epsilon)),$$

where  $\mathcal{U}_{\delta}(\mathcal{M}) = \{z \in \mathbb{C}; |z - s| < \delta \text{ for some } s \in \mathcal{M}\}$  is a  $\delta$ -neighborhood of  $\mathcal{M} \subset \mathbb{C}$ .

*Example 1.1.* The relation (1.7) is illustrated in Figure 1.1. The underlying norm in the following explanation is the spectral norm.

The upper row of the figure deals with the case  $m = 1$ . The first two pictures show the sets  $\mathcal{C}_{\mathbb{R}^{3 \times 3}}(A, \lambda, \epsilon)$  for the matrix

$$A = \begin{bmatrix} -3 & -10 & -10 \\ 1 & 5 & 5 \\ 0 & -2 & -2 \end{bmatrix}$$

and its simple eigenvalue  $\lambda = i$ . A corresponding pair of right and left eigenvectors satisfying  $y^*x = 1$  is given by

$$x = [1 + 3i \quad -2 - i \quad 2]^\top, \quad y = (1/2)[1 \quad 3 - i \quad 3 - i]^\top.$$

The right picture in the upper row shows the set  $K_{\mathbb{R}^{n \times n}}^{(1)}(x, y) = K_{\mathbb{R}^{n \times n}}(x, y)$ . By (1.7) we have

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{C}_{\mathbb{R}^{3 \times 3}}(A, i, \epsilon) - i}{\epsilon} = K_{\mathbb{R}^{3 \times 3}}(x, y).$$

The pictures indicate the convergence. The scalings have been chosen such that the displayed sets have approximately the same size. The plots of the pseudospectra components  $\mathcal{C}_{\mathbb{R}^{3 \times 3}}(A, i, \epsilon)$  have been generated using the formula

$$\sigma_{\mathbb{R}^{n \times n}}(A, \epsilon) = \{s \in \mathbb{C}; \tilde{\tau}_n(sI - A) \leq \epsilon\}, \quad A \in \mathbb{C}^{n \times n}, \epsilon > 0.$$

Here  $\tilde{\tau}_n$  denotes the smallest real perturbation value of the second kind [1], which is given by

$$\tilde{\tau}_n(M) = \sup_{\gamma \in (0, 1]} \sigma_{2n-1} \left( \begin{bmatrix} \Re M & -\gamma \Im M \\ \gamma^{-1} \Im M & \Re M \end{bmatrix} \right), \quad M \in \mathbb{C}^{n \times n},$$

where  $\sigma_{2n-1}$  is the second smallest singular value. The set  $K_{\mathbb{R}^{3 \times 3}}(x, y)$  has been computed using Theorem 6.5 below.

The left pictures in the lower row of the figure show the real pseudospectra  $\sigma_{\mathbb{R}^{3 \times 3}}(J_3, \epsilon) = \mathcal{C}_{\mathbb{R}^{3 \times 3}}(J_3, 0, \epsilon)$  for the 3 by 3 Jordan block

$$J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The right picture shows the limit set  $K_{\mathbb{R}^{n \times n}}^{(3)}(e_1, e_3)$ , where  $e_1 = [1 \ 0 \ 0]^\top$ ,  $e_3 = [0 \ 0 \ 1]^\top$ . Note that  $e_1$  is a right eigenvector and  $e_1^*$  is a left generalized eigenvector of  $J_3$  satisfying  $e_1^*e_1 = 1$ ,  $e_1^*J_3^2 = e_3^*$ . Hence, (1.7) yields

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{C}_{\mathbb{R}^{3 \times 3}}(J_3, 0, \epsilon)}{\epsilon^{1/3}} = K_{\mathbb{R}^{3 \times 3}}^{(3)}(e_1, e_3).$$

It is easily verified that the set  $K_{\mathbb{R}^{3 \times 3}}(e_1, e_3)$  equals the interval  $[-1, 1]$ . Thus,

$$K_{\mathbb{R}^{3 \times 3}}^{(3)}(e_3, e_1) = [-1, 1] \cup e^{\pi i/3}[-1, 1] \cup e^{2\pi i/3}[-1, 1].$$

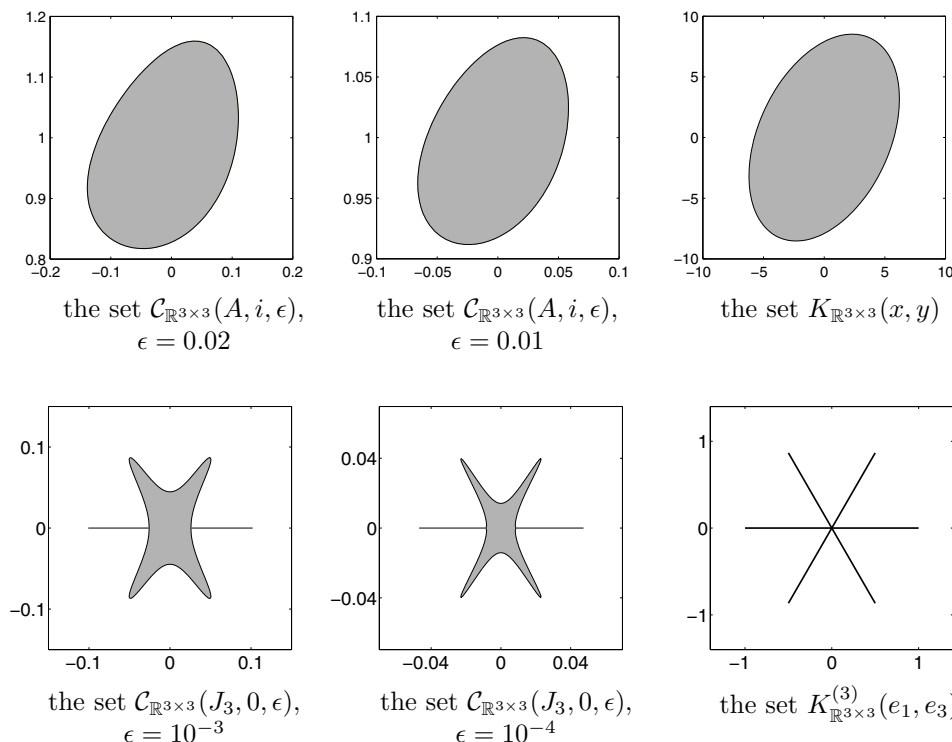


FIG. 1.1. The sets defined in Example 1.1.

The aim of this paper is to provide methods for calculating the sets  $K_{\Delta}(x, y)$ . In particular we consider the following perturbation classes  $\Delta$ :

$$\begin{aligned}
 & \mathbb{K}^{n \times n}, \\
 & \text{Sym}_{\mathbb{K}} = \{ \Delta \in \mathbb{K}^{n \times n}; \Delta^{\top} = \Delta \}, \\
 & \text{Skew}_{\mathbb{K}} = \{ \Delta \in \mathbb{K}^{n \times n}; \Delta^{\top} = -\Delta \}, \\
 & \text{Herm} = \{ \Delta \in \mathbb{C}^{n \times n}; \Delta^* = \Delta \}, \quad \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}.
 \end{aligned}
 \tag{1.8}$$

Our considerations are based on two observations concerning  $K_{\Delta}(x, y)$ ,  $\Delta \subseteq \mathbb{C}^{n \times n}$ :

(A) If  $\Delta \in \Delta$  implies that  $z\Delta \in \Delta$  for all  $z \in \mathbb{C}$ , then  $K_{\Delta}(x, y)$  is a disk. The  $m$ th root of the radius of that disk equals the condition number  $\kappa_{\Delta}(A, \lambda)$ .

(B) If  $\Delta$  is convex, then  $K_{\Delta}(x, y)$  is convex, too.

Statement (A) yields that  $K_{\Delta}(x, y)$  is a disk for  $\Delta \in \{\mathbb{C}^{n \times n}, \text{Sym}_{\mathbb{C}}, \text{Skew}_{\mathbb{C}}\}$ . Observation (B) enables us to approximate  $K_{\Delta}(x, y)$  using its support function.

The organization of this paper is as follows. In section 2 we recall some basic facts about convex sets and support functions and specialize them to the sets  $K_{\Delta}(x, y)$ . In section 3 we characterize the support function of  $K_{\Delta}(x, y)$  for the sets  $\Delta$  in (1.8) via dual norms and orthogonal projectors. The results are then applied to the cases where the underlying norm is of Hölder type (section 4) or unitarily invariant (section 5). In section 4 we also treat zero-structured perturbations. Condition numbers for these perturbation classes with respect to the Frobenius norm have been considered by Noschese and Pasquini [17]. Our approach yields an extension of their results to other norms. Section 6 deals with  $K_{\Delta}(x, y)$  for the case where the underlying norm

is either the spectral norm or the Frobenius norm. It is shown that  $K_{\Delta}(x, y)$  is an ellipse for many important perturbation classes. The results obtained so far will be extended in section 7 to classes of matrices that are self- or skew-adjoint with respect to an inner product.

**2. Characterization by support functions.** Let  $K$  be a nonempty compact convex subset of  $\mathbb{C}$ . Then its support function  $s_K : \mathbb{C} \rightarrow \mathbb{R}$  is defined by

$$(2.1) \quad s_K(z) = \max_{\xi \in K} \Re(\bar{z}\xi) = \max_{\xi \in K} z^{\top} \xi,$$

where in the second equation the complex numbers  $z = z_1 + iz_2$ ,  $\xi = \xi_1 + i\xi_2$  have been identified with the corresponding vectors  $[z_1, z_2]^{\top}$ ,  $[\xi_1, \xi_2]^{\top} \in \mathbb{R}^2$ . The set  $K$  is uniquely determined by its support function since we have [9, Corollary 3.1.2]

$$(2.2) \quad K = \{\xi \in \mathbb{C}; \Re(\bar{z}\xi) \leq s_K(z) \text{ for all } z \in \mathbb{C} \text{ with } |z| = 1\}.$$

Furthermore, the boundary of  $K$  is given as

$$(2.3) \quad \partial K = \{\xi \in K; \Re(\bar{z}\xi) = s_K(z) \text{ for some } z \in \mathbb{C} \text{ with } |z| = 1\}.$$

This follows from (2.2) and the compactness of the unit circle. Let  $r_K = \max\{|\xi|; \xi \in K\}$ . Then  $r_K$  is the radius of the smallest disk about 0 that contains  $K$ . It is easily seen that  $r_K = \max\{s_K(z); z \in \mathbb{C}, |z| = 1\}$ . If  $s_K(z) = r|z|$  for some  $r \geq 0$ , then  $K$  is a disk about 0 with radius  $r = r_K$ . We will also need the following fact.

**PROPOSITION 2.1.** *Assume the nonempty compact convex set  $K \subset \mathbb{C}$  is point symmetric with respect to 0; i.e.,  $\xi \in K$  implies  $-\xi \in K$ . Assume further that  $s_K(z) = 0$  for some  $z \in \mathbb{C}$  with  $|z| = 1$ . Then  $K$  is a line segment. Specifically,*

$$K = \{\theta iz; \theta \in \mathbb{R}, |\theta| \leq s_K(iz)\}.$$

*Proof.* From the point symmetry it follows that  $s_K(z) = s_K(-z)$ . Hence, if  $s_K(z) = 0$ , then  $\Re(\bar{z}\xi) = 0$  for all  $\xi \in K$ . Thus  $K \subset \{\theta iz; \theta \in \mathbb{R}\}$ . By compactness and convexity,  $K = \{\theta iz; \theta \in \mathbb{R}, |\theta| \leq r\}$  for some  $r \geq 0$ . It is easily verified that  $r = s_K(iz)$  if  $|z| = 1$ .  $\square$

The relations (2.2) and (2.3) can be used to approximate  $K$  via the following method (adapted from the standard algorithm for approximating the boundary of the field of values [11, section 1.5]). Let  $z_j = e^{i\phi_j}$ ,  $j \in \underline{N}$ , where  $0 = \phi_1 < \phi_2 < \dots < \phi_N < 2\pi$ . Let  $\xi_j \in K$ ,  $j \in \underline{N}$ , be such that  $\Re(\bar{z}_j \xi_j) = s_K(z_j)$ . Then by (2.3) each  $\xi_j$  is a boundary point of  $K$ . Let  $K_1$  denote the convex hull of these points, and let  $K_2 = \{\xi \in \mathbb{C}; \Re(\bar{z}_j \xi) \leq s_K(z_j), j \in \underline{N}\}$ . Then we have  $K_1 \subseteq K \subseteq K_2$ , where the latter inclusion follows from (2.2). The boundary of  $K_1$  is a polygon with vertices  $\xi_1, \xi_2, \dots, \xi_N$ .

The proposition below yields the basis for our further development.

**PROPOSITION 2.2.** *Let  $\Delta$  be a nonempty convex subset of  $\mathbb{C}^{n \times n}$ . Then the following hold:*

- (i) *The set  $K_{\Delta}(x, y)$  defined in (1.1) is a compact convex subset of  $\mathbb{C}$  with support function*

$$(2.4) \quad s_{\Delta}(z) = \max_{\substack{\Delta \in \Delta \\ \|\Delta\| \leq 1}} \Re(\bar{z} y^* \Delta x) = \max_{\substack{\Delta \in \Delta \\ \|\Delta\| \leq 1}} \Re \operatorname{tr}(\Delta^*(z y x^*)), \quad z \in \mathbb{C}.$$

*If  $\Delta$  is a cone, then the maximum is attained for some  $\Delta \in \Delta$  with  $\|\Delta\| = 1$ .*

- (ii) Let  $|z| = 1$ , and let  $\Delta_z \in \Delta$  be a maximizer for (2.4). Then  $y^* \Delta_z x$  is a boundary point of  $K_\Delta(x, y)$ .
- (iii) Suppose  $\Delta$  is a vector space over  $\mathbb{R}$ , and let  $s_\Delta(z) = 0$  for some  $z \in \mathbb{C}$  with  $|z| = 1$ . Then  $K_\Delta(x, y)$  is a line segment. Specifically,

$$K_\Delta(x, y) = \{ \theta iz; \theta \in \mathbb{R}, |\theta| \leq s_\Delta(iz) \}.$$

*Proof.* The compactness and convexity of  $K_\Delta(x, y)$  are obvious. Equation (2.4) is immediate from (2.1) and the relations

$$\bar{z} y^* \Delta x = \operatorname{tr}(\bar{z} y^* \Delta x) = \operatorname{tr}(\bar{z} x y^* \Delta) = \operatorname{tr}((z y x^*)^* \Delta) = \overline{\operatorname{tr}(\Delta^*(z y x^*))}.$$

(ii) follows from (2.3). (iii) is a consequence of Proposition 2.1.  $\square$

**3. Dual norms and orthogonal projectors.** The dual of a vector norm  $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$  is defined by

$$(3.1) \quad \|x\|' = \max_{\substack{y \in \mathbb{C}^n \\ \|y\| = 1}} \Re(y^* x), \quad x \in \mathbb{C}^n.$$

There is a natural extension of this definition to matrix norms.

DEFINITION 3.1. Let  $\|\cdot\|$  be a norm on  $\mathbb{C}^{n \times n}$ . Then its dual is defined as

$$(3.2) \quad \|X\|' := \max_{\substack{Y \in \mathbb{C}^{n \times n} \\ \|Y\| = 1}} \Re \operatorname{tr}(Y^* X), \quad X \in \mathbb{C}^{n \times n}.$$

This yields the following corollary to Proposition 2.2.

COROLLARY 3.2. For any norm  $\|\cdot\|$  on  $\mathbb{C}^{n \times n}$  the support function  $s_{\mathbb{C}^{n \times n}}$  of  $K_{\mathbb{C}^{n \times n}}(x, y)$  is given by  $s_{\mathbb{C}^{n \times n}}(z) = |z| \|yx^*\|'$ ,  $z \in \mathbb{C}$ . Thus  $K_{\mathbb{C}^{n \times n}}(x, y)$  is a disk of radius  $\|yx^*\|'$ .

The map

$$(3.3) \quad (X, Y) \mapsto \Re \operatorname{tr}(Y^* X)$$

is a positive definite symmetric  $\mathbb{R}$ -bilinear form on  $\mathbb{C}^{n \times n}$ . Thus for each subspace (over  $\mathbb{R}$ )  $\Delta \subseteq \mathbb{C}^{n \times n}$  we have the direct decomposition  $\mathbb{C}^{n \times n} = \Delta \oplus \Delta^\perp$ , where  $\Delta^\perp = \{X \in \mathbb{C}^{n \times n}; \Re \operatorname{tr}(\Delta^* X) = 0 \text{ for all } \Delta \in \Delta\}$  is the orthogonal complement of  $\Delta$  with respect to the inner product (3.3). The orthogonal projector onto  $\Delta$  is the  $\mathbb{R}$ -linear map  $\mathcal{P}_\Delta : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$  satisfying

$$\mathcal{P}_\Delta(X_1 + X_2) = X_1 \quad \text{for all } X_1 \in \Delta, X_2 \in \Delta^\perp.$$

Note that for all  $X, Y \in \mathbb{C}^{n \times n}$ ,

$$(3.4) \quad \Re \operatorname{tr}(\mathcal{P}_\Delta(Y)^* X) = \Re \operatorname{tr}(\mathcal{P}_\Delta(Y)^* \mathcal{P}_\Delta(X)) = \Re \operatorname{tr}(Y^* \mathcal{P}_\Delta(X)).$$

We will need the following lemma.

LEMMA 3.3. Let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a vector space over  $\mathbb{C}$ , and let  $\Delta_{\mathbb{R}} = \Delta \cap \mathbb{R}^{n \times n}$ . Suppose that  $\Delta = \Delta_{\mathbb{R}} \oplus i\Delta_{\mathbb{R}}$ . Then the orthogonal projector onto  $\Delta_{\mathbb{R}}$  satisfies

$$(3.5) \quad \mathcal{P}_{\Delta_{\mathbb{R}}}(X) = \Re(\mathcal{P}_\Delta(X)) \quad \text{for all } X \in \mathbb{C}^{n \times n}.$$

*Proof.* This is immediate from the fact that the vector spaces  $\Delta_{\mathbb{R}}$  and  $i\Delta_{\mathbb{R}}$  are orthogonal to each other with respect to the inner product (3.3).  $\square$

The table below gives the orthogonal projectors for the subspaces introduced in (1.8):

$\Delta$	$\mathcal{P}_\Delta(X)$
$\mathbb{C}^{n \times n}$	$X$
$\mathbb{R}^{n \times n}$	$\Re X$
Herm	$(X + X^*)/2$
Sym $_{\mathbb{C}}$	$(X + X^\top)/2$
Skew $_{\mathbb{C}}$	$(X - X^\top)/2$
Sym $_{\mathbb{R}}$	$\Re(X + X^\top)/2$
Skew $_{\mathbb{R}}$	$\Re(X - X^\top)/2$

The main results of this paper are based on the next lemma.

LEMMA 3.4. *Let  $\|\cdot\|$  be a norm on  $\mathbb{C}^{n \times n}$ , and let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a vector space over  $\mathbb{R}$ . Suppose the orthogonal projector onto  $\Delta$  is a contraction with respect to the norm  $\|\cdot\|$ , i.e.,*

$$(3.7) \quad \|\mathcal{P}_\Delta(X)\| \leq \|X\| \quad \text{for all } X \in \mathbb{C}^{n \times n}.$$

Then, for all  $M \in \mathbb{C}^{n \times n}$ ,

$$(3.8) \quad \max_{\substack{\Delta \in \Delta \\ \|\Delta\| = 1}} \Re \operatorname{tr}(\Delta^* M) = \|\mathcal{P}_\Delta(M)\|'.$$

Let  $\Delta_0 \in \mathbb{C}^{n \times n}$  be such that  $\|\Delta_0\| = 1$ , and let  $\Re \operatorname{tr}(\Delta_0^* \mathcal{P}_\Delta(M)) = \|\mathcal{P}_\Delta(M)\|'$ . If  $\mathcal{P}_\Delta(M) \neq 0$ , then the matrix  $\Delta_1 = \mathcal{P}_\Delta(\Delta_0)$  is a maximizer for the left-hand side of (3.8).

*Proof.* Let  $L$  denote the left-hand side of (3.8). For  $\Delta \in \Delta$  we have  $\Re \operatorname{tr}(\Delta^* M) = \Re \operatorname{tr}(\Delta^* \mathcal{P}_\Delta(M))$ . This yields  $L \leq \|\mathcal{P}_\Delta(M)\|'$ . We show the opposite inequality. For the matrix  $\Delta_0$  we have  $\|\mathcal{P}_\Delta(M)\|' = \Re \operatorname{tr}(\Delta_0^* \mathcal{P}_\Delta(M)) = \Re \operatorname{tr}(\mathcal{P}_\Delta(\Delta_0)^* \mathcal{P}_\Delta(M))$ . If  $\mathcal{P}_\Delta(\Delta_0) = 0$ , then  $\|\mathcal{P}_\Delta(M)\|' = 0 = L$ . Suppose  $\mathcal{P}_\Delta(\Delta_0) \neq 0$ . By condition (3.7) we have  $\|\mathcal{P}_\Delta(\Delta_0)\| \leq \|\Delta_0\| = 1$ . The matrix  $\Delta_1 = \mathcal{P}_\Delta(\Delta_0)/\|\mathcal{P}_\Delta(\Delta_0)\|$  satisfies  $\|\Delta_1\| = 1$  and  $\Re \operatorname{tr}(\Delta_1^* \mathcal{P}_\Delta(M)) = \|\mathcal{P}_\Delta(M)\|'/\|\mathcal{P}_\Delta(\Delta_0)\| \geq \|\mathcal{P}_\Delta(M)\|'$ . Thus  $L \geq \|\mathcal{P}_\Delta(M)\|'$ . Consequently,  $L = \|\mathcal{P}_\Delta(M)\|'$  and  $\|\mathcal{P}_\Delta(\Delta_0)\| = 1$ .  $\square$

Remark 3.5. The Frobenius norm  $\|X\|_F = \sqrt{\operatorname{tr}(X^* X)}$  is the norm induced by the inner product  $(X, Y) \mapsto \Re \operatorname{tr}(X^* Y)$ . Hence, the Frobenius norm is its own dual, and an orthogonal projection onto a subspace (over  $\mathbb{R}$  or  $\mathbb{C}$ ) is always a contraction with respect to the Frobenius norm. Thus, all statements of Lemma 3.4 hold for  $\|\cdot\| = \|\cdot\|_F$ , and we have  $\|\mathcal{P}_\Delta(M)\|'_F = \|\mathcal{P}_\Delta(M)\|_F$ .

From Proposition 2.2 and Lemma 3.4 (applied to the matrix  $M = z y x^*$ ,  $z \in \mathbb{C}$ ) we obtain the following theorem.

THEOREM 3.6. *Let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a vector space over  $\mathbb{R}$ , and let  $s_\Delta : \mathbb{C} \rightarrow \mathbb{R}$  denote the support function of  $K_\Delta(x, y)$ . Suppose (3.7) holds for the underlying norm. Then the following hold:*

(i) *The support function satisfies*

$$(3.9) \quad s_\Delta(z) = \|\mathcal{P}_\Delta(z y x^*)\|', \quad z \in \mathbb{C}.$$

- (ii) Let  $|z| = 1$ , and let  $\Delta_z \in \mathbb{C}^{n \times n}$  be such that  $\|\Delta_z\| = 1$  and  $\Re \operatorname{tr}(\Delta_z^* \mathcal{P}_\Delta(z y x^*)) = s_\Delta(z)$ . Then  $y^* \mathcal{P}_\Delta(\Delta_z) x \in \mathbb{C}$  is a boundary point of  $K_\Delta(x, y)$ . If  $y^* \mathcal{P}_\Delta(\Delta_z) x = 0$ , then  $K_\Delta(x, y)$  is a line segment.
- (iii) If  $\Delta$  is a vector space over  $\mathbb{C}$ , then

$$(3.10) \quad s_\Delta(z) = \|\mathcal{P}_\Delta(y x^*)\|' |z|, \quad z \in \mathbb{C}.$$

Thus  $K_\Delta(x, y)$  is a disk about 0 with radius  $\|\mathcal{P}_\Delta(y x^*)\|'$ .

Next, we consider norms that have one of the following properties (a)–(c) for all  $X \in \mathbb{C}^{n \times n}$ :

$$(3.11) \quad (a) \|X\| = \|\bar{X}\|, \quad (b) \|X\| = \|X^*\|, \quad (c) \|X\| = \|X^\top\|.$$

Note that two of these conditions imply the third.

LEMMA 3.7. Let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a vector space over  $\mathbb{C}$  such that  $\Delta = \Delta_{\mathbb{R}} \oplus i\Delta_{\mathbb{R}}$ , where  $\Delta_{\mathbb{R}} = \Delta \cap \mathbb{R}^{n \times n}$ . Let  $\|\cdot\|$  be a norm such that conditions (a) and (3.7) hold. Then condition (3.7) also holds for the orthogonal projector onto  $\Delta_{\mathbb{R}}$ ; i.e., we have  $\|\mathcal{P}_{\Delta_{\mathbb{R}}}(X)\| \leq \|X\|$  for all  $X \in \mathbb{C}^{n \times n}$ .

*Proof.* We have, for  $X \in \mathbb{C}^{n \times n}$ ,

$$\begin{aligned} \|\mathcal{P}_{\Delta_{\mathbb{R}}}(X)\| &= \|\Re(\mathcal{P}_\Delta(X))\| && \text{(by Lemma 3.3)} \\ &= \frac{1}{2} \|\mathcal{P}_\Delta(X) + \overline{\mathcal{P}_\Delta(X)}\| \\ &\leq \frac{1}{2} (\|\mathcal{P}_\Delta(X)\| + \|\overline{\mathcal{P}_\Delta(X)}\|) \\ &= \|\mathcal{P}_\Delta(X)\| && \text{(by condition (a))} \\ &\leq \|X\|. \quad \square \end{aligned}$$

LEMMA 3.8. Condition (3.7) holds for the following cases:

- (i) The norm satisfies (a) and  $\Delta = \mathbb{R}^{n \times n}$ .  
(ii) The norm satisfies (b) and  $\Delta = \operatorname{Herm}$ .  
(iii) The norm satisfies (c) and  $\Delta \in \{\operatorname{Sym}_{\mathbb{C}}, \operatorname{Skew}_{\mathbb{C}}\}$ .  
(iv) The norm satisfies (a), (b), and (c) and  $\Delta \in \{\operatorname{Sym}_{\mathbb{R}}, \operatorname{Skew}_{\mathbb{R}}\}$ .

*Proof.* (i) is immediate from Lemma 3.7. (b) implies that  $\|\mathcal{P}_{\operatorname{Herm}}(X)\| = \|(X + X^*)/2\| \leq (\|X\| + \|X^*\|)/2 = \|X\|$ . This yields (ii). The proof of (iii) is analogous. (iv) follows from (iii) and Lemma 3.7.  $\square$

THEOREM 3.9. The following assertions hold for the support function  $s_\Delta : \mathbb{C} \rightarrow \mathbb{R}$  of  $K_\Delta(x, y)$ :

- (i) If the norm  $\|\cdot\|$  satisfies condition (a), then  $s_{\mathbb{R}^{n \times n}}(z) = \|\Re(z y x^*)\|'$ .  
(ii) If the norm  $\|\cdot\|$  satisfies condition (b), then  $s_{\operatorname{Herm}}(z) = \|\mathcal{P}_{\operatorname{Herm}}(z y x^*)\|'$ .  
(iii) If the norm  $\|\cdot\|$  satisfies condition (c), then

$$s_{\operatorname{Sym}_{\mathbb{C}}}(z) = \|\mathcal{P}_{\operatorname{Sym}_{\mathbb{C}}}(z y x^*)\|' = \|\mathcal{P}_{\operatorname{Sym}_{\mathbb{C}}}(y x^*)\|' |z|$$

and

$$s_{\operatorname{Skew}_{\mathbb{C}}}(z) = \|\mathcal{P}_{\operatorname{Skew}_{\mathbb{C}}}(z y x^*)\|' = \|\mathcal{P}_{\operatorname{Skew}_{\mathbb{C}}}(y x^*)\|' |z|.$$

- (iv) If the norm  $\|\cdot\|$  satisfies (a), (b), and (c), then

$$s_{\operatorname{Sym}_{\mathbb{R}}}(z) = \|\mathcal{P}_{\operatorname{Sym}_{\mathbb{R}}}(z y x^*)\|' \quad \text{and} \quad s_{\operatorname{Skew}_{\mathbb{R}}}(z) = \|\mathcal{P}_{\operatorname{Skew}_{\mathbb{R}}}(z y x^*)\|'.$$

In the next sections we specialize Theorem 3.9 to classes of norms for which the duals can be explicitly given.



**4. Norms of Hölder type.** The Hölder- $p$ -norm of  $x = [x_1, \dots, x_n]^\top \in \mathbb{C}^n$  is defined by

$$(4.1) \quad \|x\|_p = \begin{cases} \left( \sum_{j \in \underline{n}} |x_j|^p \right)^{1/p} & \text{for } 1 \leq p < \infty, \\ \max_{j \in \underline{n}} |x_j| & \text{for } p = \infty. \end{cases}$$

We consider the matrix norms of Hölder type [8, page 717] defined by

$$(4.2) \quad \|X\|_{r|p} = \left\| \left[ \|x_1^\top\|_r, \dots, \|x_n^\top\|_r \right]^\top \right\|_p, \quad 1 \leq p, r \leq \infty,$$

where  $x_1, \dots, x_n$  denote the rows of  $X \in \mathbb{C}^{n \times n}$ . Note that  $\|X\|_{1|\infty}$  is the row sum norm and

$$(4.3) \quad \|X\|_{p|p} = \begin{cases} \left( \sum_{j,k \in \underline{n}} |x_{jk}|^p \right)^{1/p} & \text{for } 1 \leq p < \infty, \\ \max_{j,k \in \underline{n}} |x_{jk}| & \text{for } p = \infty, \end{cases}$$

where the numbers  $x_{jk}$  are the entries of  $X$ . In particular,  $\|\cdot\|_{2|2}$  is the Frobenius norm.

As is well known the dual of the Hölder- $p$ -norm is the Hölder- $q$ -norm, where  $\frac{1}{p} + \frac{1}{q} = 1$  if  $1 \leq p < \infty$  and  $q = 1$  if  $p = \infty$ . Using this fact the next proposition is easily verified.

**PROPOSITION 4.1.** *The dual of the norm  $\|\cdot\|_{r|p}$  is  $\|\cdot\|_{t|q}$ , where*

$$(4.4) \quad \begin{aligned} \frac{1}{r} + \frac{1}{t} &= 1 \text{ if } 1 \leq r < \infty \text{ and } t = 1 \text{ if } r = \infty, \\ \frac{1}{p} + \frac{1}{q} &= 1 \text{ if } 1 \leq p < \infty \text{ and } q = 1 \text{ if } p = \infty. \end{aligned}$$

To a given  $X \in \mathbb{C}^{n \times n}$  with rows  $x_1, \dots, x_n$  a matrix  $Y_0 \in \mathbb{C}^{n \times n}$  satisfying  $\|Y_0\|_{t|q} = 1$  and  $\Re \operatorname{tr}(Y_0^* X) = \Re \operatorname{tr}(XY_0^*) = \|X\|_{r|p}$  can be constructed via the following procedure. Let  $\xi = [\|x_1^\top\|_r, \dots, \|x_n^\top\|_r]^\top$ . Choose a nonnegative vector  $\eta = [\eta_1, \dots, \eta_n]^\top$  such that  $\|\eta\|_q = 1$  and  $\eta^\top \xi = \|\xi\|_p$ . For each  $j \in \underline{n}$  choose a row vector  $y_j \in \mathbb{C}^n$  with  $\|y_j^\top\|_t = \eta_j$  and  $x_j y_j^\top = \|x_j^\top\|_r \eta_j$ . Then  $Y_0 = [y_1^\top, \dots, y_n^\top]^*$  has the required properties.

From Proposition 4.1 combined with Corollary 3.2 and Theorem 3.9 we get the following corollary.

**COROLLARY 4.2.** *Let  $1 \leq r, p \leq \infty$ , and let  $t, q$  be given by (4.4). Let  $K_\Delta(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\|_{r|p} \leq 1\}$ . Then the following hold:*

- (i) *the set  $K_{\mathbb{C}^{n \times n}}(x, y)$  is a disk of radius  $\|yx^*\|_{t|q}$ ;*
- (ii) *the support function of  $K_{\mathbb{R}^{n \times n}}(x, y)$  is*

$$s_{\mathbb{R}^{n \times n}}(z) = \|\Re(z y x^*)\|_{t|q}, \quad z \in \mathbb{C};$$

- (iii) *for the case  $p = r$  and  $\Delta \in \{\text{Herm}, \text{Sym}_{\mathbb{C}}, \text{Skew}_{\mathbb{C}}, \text{Sym}_{\mathbb{R}}, \text{Skew}_{\mathbb{R}}\}$  the support function of  $K_\Delta(x, y)$  is  $s_\Delta(z) = \|\mathcal{P}_\Delta(z y x^*)\|_{q|q}$ .*

**Example 4.3.** Figure 4.1 shows the sets

$$(4.5) \quad \begin{aligned} K_{\mathbb{R}^{n \times n}}(x, y) &= \{y^* \Delta x; \Delta \in \mathbb{R}^{n \times n}, \|\Delta\|_{1|\infty} \leq 1\}, \\ K_{\mathbb{R}^{n \times n}}^{(3)}(x, y) &= \{z \in \mathbb{C}; z^3 \in K_{\mathbb{R}^{n \times n}}(x, y)\}, \end{aligned}$$

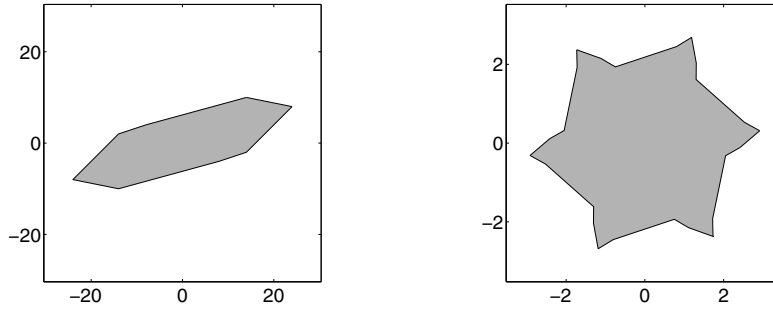


FIG. 4.1. The sets  $K_{\mathbb{R}^{n \times n}}(x, y)$  (left) and  $K_{\mathbb{R}^{n \times n}}^{(3)}(x, y)$  (right) from Example 4.3.

where

$$x = [1 - i, -4, 1]^\top, \quad y = [2 - i, 2, 2 - i]^\top.$$

The plot of  $K_{\mathbb{R}^{n \times n}}(x, y)$  has been generated by computing 200 boundary points using claim (ii) of Theorem 3.6 and Proposition 4.1.

The following theorem extends a result by Noschese and Pasquini [17] on zero-structured perturbations.

**THEOREM 4.4.** *For an index set  $\mathcal{I} \subseteq \underline{n} \times \underline{n}$ , let  $\Delta(\mathcal{I})$  denote the set of  $\Delta = [\delta_{jk}] \in \mathbb{C}^{n \times n}$  with  $\delta_{jk} = 0$  for  $(j, k) \notin \mathcal{I}$ .*

(i) *The orthogonal projection of  $X = [x_{jk}] \in \mathbb{C}^{n \times n}$  onto  $\Delta(\mathcal{I})$  is given by*

$$(4.6) \quad \mathcal{P}_{\Delta(\mathcal{I})}(X) = [\tilde{x}_{jk}], \quad \text{where } \tilde{x}_{jk} = \begin{cases} x_{jk} & \text{if } (j, k) \in \mathcal{I}, \\ 0 & \text{if } (j, k) \notin \mathcal{I}. \end{cases}$$

*The orthogonal projection onto  $\Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}$  satisfies*

$$(4.7) \quad \mathcal{P}_{\Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}}(X) = \Re(\mathcal{P}_{\Delta(\mathcal{I})}(X)) = [\hat{x}_{jk}],$$

*where*

$$\hat{x}_{jk} = \begin{cases} \Re x_{jk} & \text{if } (j, k) \in \mathcal{I}, \\ 0 & \text{if } (j, k) \notin \mathcal{I}. \end{cases}$$

(ii) *Let  $p, q \in [1, \infty]$  be such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Then the support function of  $K_{\Delta(\mathcal{I})}(x, y) = \{y^* \Delta x; \Delta \in \Delta(\mathcal{I}), \|\Delta\|_{p|p} \leq 1\}$  is given by*

$$s_{\Delta(\mathcal{I})}(z) = \|\mathcal{P}_{\Delta(\mathcal{I})}(z y x^*)\|_{q|q} = \|\mathcal{P}_{\Delta(\mathcal{I})}(y x^*)\|_{q|q} |z|, \quad z \in \mathbb{C}.$$

*Hence  $K_{\Delta(\mathcal{I})}(x, y)$  is a disk about 0 of radius  $\|\mathcal{P}_{\Delta(\mathcal{I})}(y x^*)\|_{q|q}$ . Furthermore, the support function of  $K_{\Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}}(x, y) = \{y^* \Delta x; \Delta \in \Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}, \|\Delta\|_{p|p} \leq 1\}$  is given by*

$$s_{\Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}}(z) = \|\mathcal{P}_{\Delta(\mathcal{I}) \cap \mathbb{R}^{n \times n}}(z y x^*)\|_{q|q} = \|\Re(z \mathcal{P}_{\Delta(\mathcal{I})}(y x^*))\|_{q|q}, \quad z \in \mathbb{C}.$$

*Proof.* Obviously, the orthogonal complement of  $\Delta(\mathcal{I})$  is  $\Delta((\underline{n} \times \underline{n}) \setminus \mathcal{I})$ . This yields (4.6). Equation (4.7) follows from Lemma 3.3. It is easily seen that the projectors in (i) satisfy the contraction condition (3.7) for  $\|\cdot\| = \|\cdot\|_{p|p}$ . This yields (ii).  $\square$

**5. Unitarily invariant norms.** In what follows,  $\mathcal{U}_n$  denotes the set of all unitary  $n \times n$  matrices. A norm  $\|\cdot\|$  on  $\mathbb{C}^{n \times n}$  is said to be unitarily invariant if  $\|UXV\| = \|X\|$  for all  $X \in \mathbb{C}^{n \times n}$ ,  $U, V \in \mathcal{U}_n$ . There is a one to one correspondence between the unitarily invariant norms on  $\mathbb{C}^{n \times n}$  and symmetric gauge functions [22, section II.3]. A symmetric gauge function  $\Phi$  is a symmetric and absolute norm on  $\mathbb{R}^n$ . The unitarily invariant norm  $\|\cdot\|_\Phi$  associated with  $\Phi$  is given by

$$(5.1) \quad \|X\|_\Phi := \Phi([\sigma_1(X), \sigma_2(X), \dots, \sigma_n(X)]^\top),$$

where  $\sigma_1(X) \geq \sigma_2(X) \geq \dots \geq \sigma_n(X)$  denote the singular values of  $X \in \mathbb{C}^{n \times n}$ . The unitarily invariant norm induced by the Hölder- $p$ -norm is called the Schatten- $p$ -norm, which we denote by

$$\|X\|_{(p)} := \begin{cases} \left( \sum_{k \in \mathbb{N}} \sigma_k(X)^p \right)^{1/p} & \text{if } 1 \leq p < \infty, \\ \sigma_1(X) & \text{if } p = \infty. \end{cases}$$

Note that  $\|X\|_{(\infty)}$  is the spectral norm and  $\|X\|_{(2)} = \sqrt{\text{tr}(X^*X)}$  is the Frobenius norm of  $X$ . In the following,  $\Phi'$  stands for the dual of the symmetric gauge function  $\Phi$ , i.e.,

$$(5.2) \quad \Phi'(\xi) = \max_{\substack{\eta \in \mathbb{R}^n \\ \Phi(\eta) = 1}} \eta^\top \xi, \quad \xi \in \mathbb{R}^n.$$

Let  $X = U \text{diag}(\sigma) V^*$  be a singular value decomposition, where  $U, V \in \mathcal{U}_n$  and  $\sigma = [\sigma_1, \dots, \sigma_n]^\top$  is the vector of singular values of  $X \in \mathbb{C}^{n \times n}$ . Let  $\tau = [\tau_1, \dots, \tau_n]^\top$  be a nonnegative vector such that  $\Phi(\tau) = 1$  and  $\tau^\top \sigma = \Phi'(\sigma)$ . Let  $Y_0 = U \text{diag}(\tau) V^*$ . Then

$$(5.3) \quad \|X\|'_\Phi = \max_{\substack{Y \in \mathbb{C}^{n \times n} \\ \|Y\|_\Phi = 1}} \Re \text{tr}(Y^* X) \geq \Re \text{tr}(Y_0^* X) = \tau^\top \sigma = \Phi'(\sigma) = \|X\|_\Phi.$$

It can be shown that the inequality in (5.3) is actually an equality. Hence we have the following result [2, Proposition IV.2.11].

**PROPOSITION 5.1.** *For any symmetric gauge function  $\Phi$  the dual of the unitarily invariant norm  $\|\cdot\|_\Phi$  is  $\|\cdot\|_{\Phi'}$ .*

From (5.1) it follows that unitarily invariant norms have properties (a), (b), and (c) of (3.11). Thus, by combining Theorem 3.9 and Proposition 5.1, we get the following result.

**THEOREM 5.2.** *Let  $\Phi$  be a symmetric gauge function on  $\mathbb{R}^n$ , and let  $\Delta$  be one of the sets in (3.6). Then the support function of*

$$K_\Delta(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\|_\Phi \leq 1\}, \quad x, y \in \mathbb{C}^n,$$

*is given by*

$$s_\Delta(z) = \|\mathcal{P}_\Delta(z y x^*)\|_{\Phi'} = \Phi'([\sigma_1(z), \dots, \sigma_n(z)]^\top), \quad z \in \mathbb{C},$$

*where  $\sigma_1(z), \dots, \sigma_n(z)$  denote the singular values of  $\mathcal{P}_\Delta(z y x^*)$ .*

**6. Frobenius norm and spectral norm.** The results in this section are based on the following proposition about the support function of an ellipse.

PROPOSITION 6.1. *Let  $K \subset \mathbb{C}$  be a nonempty compact convex set with support function*

$$s_K(z) = \sqrt{a|z|^2 + \Re(bz^2)}, \quad z, b \in \mathbb{C}, a \geq |b|.$$

*Then  $K$  is an ellipse (which may degenerate to a line segment). Specifically,*

$$(6.1) \quad K = \left\{ e^{-i\phi/2} \left( \sqrt{a+|b|} \xi_1 + \sqrt{a-|b|} \xi_2 i \right); \quad \xi_1, \xi_2 \in \mathbb{R}, \xi_1^2 + \xi_2^2 \leq 1 \right\},$$

where  $\phi = \arg(b)$ .

*Proof.* Let  $E$  denote the set on the right-hand side of (6.1), and let  $s_E$  denote its support function. Let

$$\alpha = \frac{1}{2}(\sqrt{a+|b|} + \sqrt{a-|b|}) e^{-i\phi/2}, \quad \beta = \frac{1}{2}(\sqrt{a+|b|} - \sqrt{a-|b|}) e^{-i\phi/2}.$$

Then, for  $\xi_1, \xi_2 \in \mathbb{R}$ ,

$$e^{-i\phi/2} \left( \sqrt{a+|b|} \xi_1 + \sqrt{a-|b|} \xi_2 i \right) = \alpha \xi + \beta \bar{\xi}, \quad \text{where } \xi = \xi_1 + \xi_2 i \in \mathbb{C}.$$

Thus

$$\begin{aligned} s_E(z) &= \max_{|\xi| \leq 1} \Re(\bar{z}(\alpha \xi + \beta \bar{\xi})) \\ &= \max_{|\xi| \leq 1} \Re((\alpha \bar{z} + \bar{\beta} z) \xi) \\ &= |\alpha \bar{z} + \bar{\beta} z| \\ &= \sqrt{(|\alpha|^2 + |\bar{\beta}|^2)|z|^2 + 2\Re(z^2 \bar{\alpha} \bar{\beta})} \\ &= \sqrt{a|z|^2 + \Re(bz^2)}. \end{aligned}$$

Thus  $s_E = s_K$ , and consequently  $E = K$ .  $\square$

Note that the set (6.1) is a disk if  $b = 0$  and  $a > 0$ . It is a line segment if  $a = |b| > 0$ .

The next theorem characterizes the sets  $K_{\Delta}(x, y)$  if the underlying norm is the Frobenius norm  $\|\cdot\|_F$ .

THEOREM 6.2. *Let  $x, y \in \mathbb{C}^n$ , and let  $P = \mathcal{P}_{\Delta}(yx^*)$ . Furthermore, let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a vector space over  $\mathbb{C}$ . Then the following hold:*

- (i) *The set  $K_{\Delta}(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\|_F \leq 1\}$  is a disk about 0 of radius  $\|P\|_F$ .*
- (ii) *Suppose that  $\Delta = \Delta_{\mathbb{R}} \oplus i \Delta_{\mathbb{R}}$ , where  $\Delta_{\mathbb{R}} = \Delta \cap \mathbb{R}^{n \times n}$ . Then the set  $K_{\Delta_{\mathbb{R}}}(x, y) = \{y^* \Delta x; \Delta \in \Delta_{\mathbb{R}}, \|\Delta\|_F \leq 1\}$  equals the ellipse defined in (6.1) with*

$$a = \frac{1}{2} \operatorname{tr}(P^* P) = \frac{1}{2} \|P\|_F^2, \quad b = \frac{1}{2} \operatorname{tr}(P^{\top} P).$$

*Proof.* (i) is a special case of claim (iii) of Theorem 3.6.

(ii) The support function of  $K_{\Delta_{\mathbb{R}}}(x, y)$  satisfies

$$\begin{aligned}
 s_{\Delta_{\mathbb{R}}}(z) &= \|\mathcal{P}_{\Delta_{\mathbb{R}}}(z y x^*)\|_{(2)} && \text{(by claim (i) of Theorem 3.9)} \\
 &= \|\Re(zP)\|_{(2)} && \text{(by Lemma 3.3)} \\
 &= \frac{1}{2} \|zP + \overline{zP}\|_{(2)} \\
 &= \frac{1}{2} \sqrt{\operatorname{tr}((zP + \overline{zP})^*(zP + \overline{zP}))} \\
 &= \frac{1}{2} \left( |z|^2 \underbrace{(\operatorname{tr}(P^*P) + \operatorname{tr}(\bar{P}^*\bar{P}))}_{=2\operatorname{tr}(P^*P)} + \underbrace{z^2\operatorname{tr}(\bar{P}^*P) + \bar{z}^2\operatorname{tr}(P^*\bar{P})}_{=2\Re(z^2 \operatorname{tr}(P^\top P))} \right)^{1/2} \\
 &= \sqrt{a|z|^2 + \Re(bz^2)}.
 \end{aligned}$$

Now, (ii) follows from Proposition 6.1.  $\square$

Theorem 6.2 has the following consequences for structured eigenvalue condition numbers.

**COROLLARY 6.3.** *Let  $\lambda \in \mathbb{C}$  be a simple eigenvalue of  $A \in \mathbb{C}^{n \times n}$  with right eigenvector  $x \in \mathbb{C}^n$  and left eigenvector  $y \in \mathbb{C}^n$  such that  $y^*x = 1$ . Then the following statements hold for the structured condition number*

$$(6.2) \quad \kappa_{\Delta}(A, \lambda) = \max\{ |y^* \Delta x|; \Delta \in \Delta, \|\Delta\| \leq 1 \}.$$

(i) *If the set  $K_{\Delta}(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\| \leq 1\}$  equals the ellipse in (6.1), then*

$$\kappa_{\Delta}(A, \lambda) = \sqrt{a + |b|}.$$

(ii) *Suppose the norm in (6.2) is the Frobenius norm. Let  $\Delta \subseteq \mathbb{C}^{n \times n}$  be a subspace over  $\mathbb{C}$ , and let  $P = \mathcal{P}_{\Delta}(y x^*)$  denote the orthogonal projection of  $y x^*$  onto  $\Delta$ . Then*

$$(6.3) \quad \kappa_{\Delta}(A, \lambda) = \|P\|_F.$$

*Suppose additionally that  $\Delta = \Delta_{\mathbb{R}} \oplus i \Delta_{\mathbb{R}}$ , where  $\Delta_{\mathbb{R}} = \Delta \cap \mathbb{R}^{n \times n}$ . Then*

$$(6.4) \quad \kappa_{\Delta_{\mathbb{R}}}(A, \lambda) = \sqrt{\frac{1}{2} \|P\|_F^2 + \frac{1}{2} |\operatorname{tr}(P^T P)|}.$$

*In particular, the following inequality holds:*

$$\kappa_{\Delta_{\mathbb{R}}}(A, \lambda) \geq \frac{1}{\sqrt{2}} \kappa_{\Delta}(A, \lambda).$$

*Proof.* (i) is obvious and (ii) is a consequence of (i) and Theorem 6.2.  $\square$

**Remark 6.4.** Relation (6.3) is Tisseur's formula [23, section 4] in another notation. From (6.3) one obtains the results of Noschese and Pasquini [18] on the structured condition number of Toeplitz matrices by observing that the orthogonal projection of a matrix  $X \in \mathbb{C}^{n \times n}$  onto the set of Toeplitz matrices is given by replacing the entries in each diagonal of  $X$  by their arithmetic mean.

Relation (6.4) generalizes a result by Byers and Kressner [4] on condition numbers with respect to real perturbations.

The next theorem gives explicit formulae for the constants  $a$  and  $b$  if  $\Delta$  is one of the perturbation classes defined in (1.8). Here we also consider the case where the underlying norm is the spectral norm.

**THEOREM 6.5.** *Let  $\|\Delta\|$  denote either the Frobenius norm or the spectral norm of  $\Delta \in \mathbb{C}^{n \times n}$ . Let  $\Delta \subseteq \mathbb{C}^{n \times n}$ , and let  $a \geq 0$ ,  $b \in \mathbb{C}$  be as in the tables below. Then the support function of*

$$K_{\Delta}(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\| \leq 1\}$$

is given by

$$(6.5) \quad s_{\Delta}(z) = \sqrt{a|z|^2 + \Re(bz^2)}, \quad z \in \mathbb{C}.$$

Hence,  $K_{\Delta}(x, y)$  equals the ellipse defined in (6.1).

Table for the Frobenius norm:

$\Delta$	$a$	$b$
$\mathbb{C}^{n \times n}$	$\ x\ ^2 \ y\ ^2$	0
$\mathbb{R}^{n \times n}$	$\frac{1}{2} \ x\ ^2 \ y\ ^2$	$\frac{1}{2} (\overline{x^{\top} x}) (y^{\top} y)$
Herm	$\frac{1}{2} \ x\ ^2 \ y\ ^2$	$\frac{1}{2} (x^* y)^2$
Sym $_{\mathbb{C}}$	$\frac{1}{2} (\ x\ ^2 \ y\ ^2 +  x^{\top} y ^2)$	0
Skew $_{\mathbb{C}}$	$\frac{1}{2} (\ x\ ^2 \ y\ ^2 -  x^{\top} y ^2)$	0
Sym $_{\mathbb{R}}$	$\frac{1}{4} (\ x\ ^2 \ y\ ^2 +  x^{\top} y ^2)$	$\frac{1}{4} ((\overline{x^{\top} x}) (y^{\top} y) + (x^* y)^2)$
Skew $_{\mathbb{R}}$	$\frac{1}{4} (\ x\ ^2 \ y\ ^2 -  x^{\top} y ^2)$	$\frac{1}{4} ((\overline{x^{\top} x}) (y^{\top} y) - (x^* y)^2)$

Table for the spectral norm:

$\Delta$	$a$	$b$
$\mathbb{C}^{n \times n}$	$\ x\ ^2 \ y\ ^2$	0
$\mathbb{R}^{n \times n}$	$\frac{1}{2} \left[ \ x\ ^2 \ y\ ^2 + \sqrt{(\ x\ ^4 -  x^{\top} x ^2)(\ y\ ^4 -  y^{\top} y ^2)} \right]$	$\frac{1}{2} (\overline{x^{\top} x}) (y^{\top} y)$
Herm	$\ x\ ^2 \ y\ ^2 - \frac{1}{2}  y^* x ^2$	$\frac{1}{2} (x^* y)^2$
Sym $_{\mathbb{C}}$	$\ x\ ^2 \ y\ ^2$	0
Skew $_{\mathbb{C}}$	$\ x\ ^2 \ y\ ^2 -  x^{\top} y ^2$	0
Skew $_{\mathbb{R}}$	$\frac{1}{2} (\ x\ ^2 \ y\ ^2 -  x^{\top} y ^2 + \sqrt{\det(F^* F)})$	$\frac{1}{2} ((\overline{x^{\top} x}) (y^{\top} y) - (x^* y)^2)$

$F = \begin{bmatrix} x & \bar{x} & y & \bar{y} \end{bmatrix}$

Here and in the following,  $\|x\|$ ,  $\|y\|$  denote the Euclidean norm of  $x, y \in \mathbb{C}^n$ .

**Remark 6.6.** Theorem 6.5 makes no statement about the case where  $\Delta = \text{Sym}_{\mathbb{R}}$  and the underlying norm is the spectral norm. The associated sets  $K_{\text{Sym}_{\mathbb{R}}}(x, y)$  are in general not ellipses. Figure 6.1 gives two examples. It shows the sets  $K_{\text{Sym}_{\mathbb{R}}}(x_j, y_j)$ ,  $j = 1, 2$ , where

$$(6.6) \quad \begin{aligned} x_1 &= [2 + i, \quad 2 + i, \quad 2]^{\top}, & y_1 &= [-2, \quad -2, \quad 3i]^{\top}, \\ x_2 &= [1 + 2i, \quad i, \quad 2]^{\top}, & y_2 &= [i, \quad -2 + 2i, \quad 1 + 2i]^{\top}. \end{aligned}$$

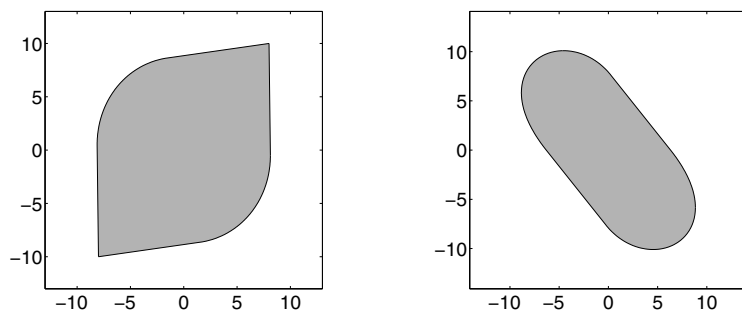


FIG. 6.1. The sets  $K_{\text{Sym}_{\mathbb{R}}}(x_1, y_1)$  (left) and  $K_{\text{Sym}_{\mathbb{R}}}(x_2, y_2)$  (right) from Remark 6.6.

According to Theorem 3.6 the boundary points of  $K_{\text{Sym}_{\mathbb{R}}}(x_j, y_j)$  are of the form  $y_j^* \mathcal{P}_{\text{Sym}_{\mathbb{R}}}(\Delta_z) x_j$ , where  $z \in \mathbb{C}$ ,  $|z| = 1$ , and  $\Delta_z \in \mathbb{C}^{n \times n}$  is any matrix satisfying  $\|\Delta_z\|_{(\infty)} = 1$  and  $\Re \text{tr}(\Delta_z^* \mathcal{P}_{\text{Sym}_{\mathbb{R}}}(z y x^*)) = \|\mathcal{P}_{\text{Sym}_{\mathbb{R}}}(z y x^*)\|_{(1)}$ . A matrix  $\Delta_z$  with these properties is given by  $\Delta_z = U_z V_z^*$ , where  $\mathcal{P}_{\text{Sym}_{\mathbb{R}}}(z y x^*) = U_z \Sigma_z V_z^*$  is a singular value decomposition of  $\mathcal{P}_{\text{Sym}_{\mathbb{R}}}(z y x^*)$ . In order to generate the plots in Figure 6.1 we have chosen 400 equidistant values of  $z$ .

The proof of the formulae for  $a$  and  $b$  with respect to the structures  $\Delta$  in the table for the spectral norm is based on the fact that for these structures the projection  $\mathcal{P}_{\Delta}(z y x^*)$  has at most two distinct nonzero singular values. However, the latter generically does not hold for the structure  $\Delta = \text{Sym}_{\mathbb{R}}$ .

The proof of Theorem 6.5 uses the lemma below.

LEMMA 6.7. Let  $M = a_1 b_1^* + a_2 b_2^*$ , where  $a_1, a_2, b_1, b_2 \in \mathbb{C}^n$ . Then the Frobenius norm and the Schatten-1-norm of  $M$  are given by

$$\begin{aligned} \|M\|_F^2 &= \|a_1\|^2 \|b_1\|^2 + \|a_2\|^2 \|b_2\|^2 + 2 \Re[(\overline{a_1^* a_2})(b_1^* b_2)], \\ \|M\|_{(1)}^2 &= \|M\|_F^2 + 2 \sqrt{(\|a_1\|^2 \|a_2\|^2 - |a_1^* a_2|^2)(\|b_1\|^2 \|b_2\|^2 - |b_1^* b_2|^2)}. \end{aligned}$$

The Frobenius norms of the matrices  $S_{\pm} = \frac{1}{2}(M \pm M^{\top})$  are given by

$$\begin{aligned} \|S_{\pm}\|_F^2 &= \frac{1}{2} (\|a_1\|^2 \|b_1\|^2 + \|a_2\|^2 \|b_2\|^2 \pm |a_1^{\top} b_1|^2 \pm |a_2^{\top} b_2|^2) \\ &\quad + \Re \left( (\overline{a_1^* a_2})(b_1^* b_2) \pm (\overline{a_1^{\top} b_2})(a_2^{\top} b_1) \right). \end{aligned}$$

The Schatten-1-norm of  $S_{-}$  satisfies

$$\|S_{-}\|_{(1)}^2 = 2 \left( \|S_{-}\|_F^2 + \sqrt{\det(A^* A)} \right),$$

where  $A = \begin{bmatrix} a_1 & a_2 & \bar{b}_1 & \bar{b}_2 \end{bmatrix} \in \mathbb{C}^{n \times 4}$ .

*Proof.* See the appendix.  $\square$

*Proof of Theorem 6.5.* First, we treat the case where  $\Delta = \mathbb{R}^{n \times n}$ . Let

$$M_z = 2 \Re(z y x^*) = z y x^* + \bar{z} \bar{y} \bar{x}^*.$$

According to Proposition 5.1 the dual of the spectral norm is the Schatten-1-norm. Hence, by Theorem 5.2 the support function of  $K_{\mathbb{R}^{n \times n}}(x, y)$  (with respect to spectral norm) is

$$\begin{aligned} s_{\mathbb{R}^{n \times n}}(z) &= \|\mathcal{P}_{\mathbb{R}^{n \times n}}(z y x^*)\|_{(1)} \quad (\text{by Theorem 5.2}) \\ &= \|\Re(z y x^*)\|_{(1)} \\ &= \frac{1}{2} \|M_z\|_{(1)} \\ &= \frac{1}{2} \sqrt{\alpha_z + 2\sqrt{\beta_z}} \quad (\text{by Lemma 6.7}), \end{aligned}$$

where

$$\begin{aligned} \alpha_z &= \|M_z\|_F^2 \\ &= \|z y\|^2 \|x\|^2 + \|\bar{z} \bar{y}\|^2 \|\bar{x}\|^2 + 2 \Re[(\overline{(z y)^* (\bar{z} \bar{y})}) (x^* \bar{x})] \\ &= 2(|z|^2 \|x\|^2 \|y\|^2 + \Re[z^2 (\overline{x^\top x}) (y^\top y)]), \\ \beta_z &= (\|z y\|^2 \|\bar{z} \bar{y}\|^2 - |(z y)^* (\bar{z} \bar{y})|^2) (\|x\|^2 \|\bar{x}\|^2 - |x^* \bar{x}|^2) \\ &= |z|^4 (\|x\|^4 - |x^\top x|^2) (\|y\|^4 - |y^\top y|^2). \end{aligned}$$

If the underlying norm is the Frobenius norm, then

$$s_{\mathbb{R}^{n \times n}}(z) = \|\Re(z y x^*)\|_F = \frac{1}{2} \|M_z\|_F = \frac{1}{2} \sqrt{\alpha_z}.$$

Next, we consider the real skew-symmetric case. Let  $S_- = \frac{1}{2}(M_z - M_z^\top)$ . The support function of  $K_{\text{Skew}_\mathbb{R}}(x, y)$  with respect to the spectral norm is

$$\begin{aligned} s_{\text{Skew}_\mathbb{R}}(z) &= \|\mathcal{P}_{\text{Skew}_\mathbb{R}}(z y x^*)\|_{(1)} \quad (\text{by Theorem 5.2}) \\ &= \frac{1}{2} \|S_-\|_{(1)} \\ &= \frac{1}{2} \sqrt{2 \|S_-\|_F^2 + 2 \sqrt{\det(A_z^* A_z)}} \quad (\text{by Lemma 6.7}) \\ &= \sqrt{\frac{1}{2} \|S_-\|_F^2 + \frac{1}{2} \sqrt{\det(A_z^* A_z)}}, \end{aligned}$$

where

$$\begin{aligned} \|S_-\|_F^2 &= \frac{1}{2} (\|z y\|^2 \|x\|^2 + \|\bar{z} \bar{y}\|^2 \|\bar{x}\|^2 - |(z y)^\top x|^2 - |(\bar{z} \bar{y})^\top \bar{x}|^2) \\ &\quad + \Re \left[ (\overline{(z y)^* (\bar{z} \bar{y})}) (x^* \bar{x}) - \overline{((z y)^\top x)} ((\bar{z} \bar{y})^\top \bar{x}) \right] \\ &= |z|^2 (\|x\|^2 \|y\|^2 - |x^\top y|^2) + \Re[z^2 (\overline{(x^\top x)} (y^\top y) - (x^* y)^2)], \\ A_z &= \begin{bmatrix} z y & \bar{z} \bar{y} & x & \bar{x} \end{bmatrix} = \underbrace{\begin{bmatrix} y & \bar{y} & x & \bar{x} \end{bmatrix}}_{=A_1} \text{diag}(z, \bar{z}, 1, 1). \end{aligned}$$



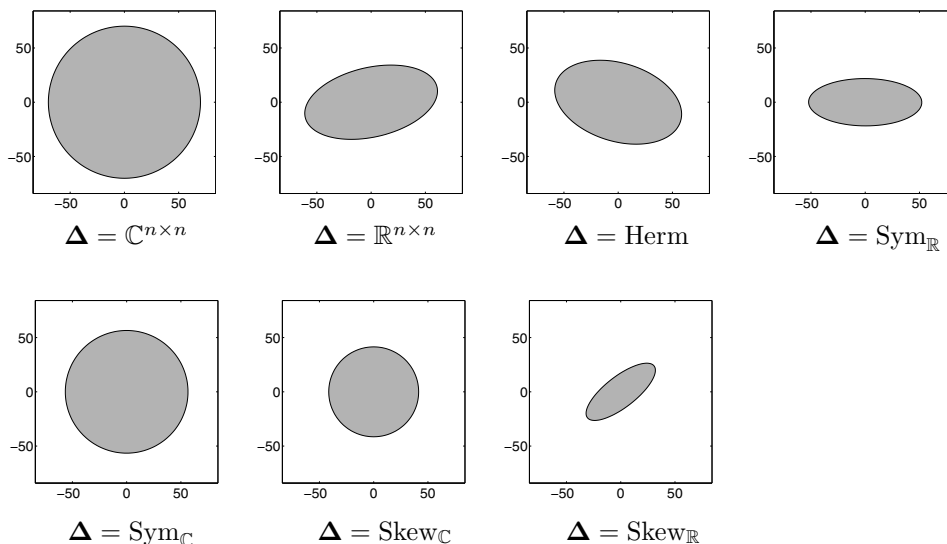


FIG. 6.2. The sets  $K_{\Delta}(x, y)$  for the Frobenius norm and  $x, y$  defined in (6.7).

We have  $\det(A_z^* A_z) = |z|^4 \det(A_1^* A_1) = |z|^4 \det(F^* F)$ , where  $F = \begin{bmatrix} x & \bar{x} & y & \bar{y} \end{bmatrix}$ .

The computations for the other cases are analogous.  $\square$

*Example 6.8.* Figure 6.2 shows the sets  $K_{\Delta}(x, y) = \{y^* \Delta x; \Delta \in \Delta, \|\Delta\|_F \leq 1\}$ , where

$$(6.7) \quad \begin{aligned} x &= [4 + 5i, 4 + 2i, 2 - i, 3 + 4i]^T, \\ y &= [3 + 3i, 1 + 3i, 5 + i, 0]^T. \end{aligned}$$

**7. Self- and skew-adjoint perturbations.** We now treat the case where  $\Delta$  is a set of matrices which are skew- or self-adjoint with respect to a scalar product on  $\mathbb{C}^n$ . Specifically we show that the associated sets  $K_{\Delta}(x, y)$  can be computed via the methods in the previous sections if the scalar product is induced by a unitary matrix and the underlying norm is unitarily invariant.

For nonsingular  $\Pi \in \mathbb{C}^{n \times n}$  we consider the scalar products

$$\langle x, y \rangle_{\Pi} = x^* \Pi y, \quad x, y \in \mathbb{C}^n, \quad \star \in \{*, \top\}.$$

Depending on whether  $\star = \top$  or  $\star = *$ , the scalar product is a bilinear form or a sesquilinear form. We assume that  $\Pi$  satisfies a symmetry relation of the form

$$(7.1) \quad \Pi^{\star} = \epsilon_0 \Pi, \quad \text{with } \epsilon_0 = -1 \text{ or } \epsilon_0 = 1.$$

A matrix  $\Delta \in \mathbb{C}^{n \times n}$  is said to be self-adjoint (skew-adjoint) with respect to the scalar product  $\langle \cdot, \cdot \rangle_{\Pi}$  if

$$(7.2) \quad \langle \Delta x, y \rangle_{\Pi} = \epsilon \langle x, \Delta y \rangle_{\Pi} \quad \text{for all } x, y \in \mathbb{C}^n,$$

and  $\epsilon = 1$  ( $\epsilon = -1$ ). The relation (7.2) is easily seen to be equivalent to

$$(7.3) \quad \Delta^* \Pi = \epsilon \Pi \Delta.$$

We denote the sets of self- and skew-adjoint matrices by

$$\text{struct}(\Pi, \star, \epsilon) := \{ \Delta \in \mathbb{C}^{n \times n} ; \Delta^* \Pi = \epsilon \Pi \Delta \}.$$

The relation (7.1) implies that (7.3) is equivalent to

$$(7.4) \quad (\Pi \Delta)^* = \epsilon_0 \epsilon \Pi \Delta.$$

We thus have the following lemma.

LEMMA 7.1. *Let  $\Pi, \Delta \in \mathbb{K}^{n \times n}$ , where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ . Suppose  $\Pi^* = \epsilon_0 \Pi$  with  $\epsilon_0 = -1$  or  $\epsilon_0 = 1$ . Then the following equivalences hold:*

$$\Delta \in \text{struct}(\Pi, \star, \epsilon) \Leftrightarrow \begin{cases} \Pi \Delta \in \text{Herm} & \text{if } \epsilon_0 \epsilon = 1, \star = *, \\ \Pi \Delta \in \text{Sym}_{\mathbb{K}} & \text{if } \epsilon_0 \epsilon = 1, \star = \top, \\ \Pi \Delta \in \text{Skew}_{\mathbb{K}} & \text{if } \epsilon_0 \epsilon = -1, \star = \top, \\ i \Pi \Delta \in \text{Herm} & \text{if } \epsilon_0 \epsilon = -1, \star = *. \end{cases}$$

In many applications  $\Pi$  is unitary. The most common examples are

$$\Pi \in \{\text{diag}(I_k, -I_{n-k}), E_n, J_n\},$$

where

$$J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \in \mathbb{C}^{2n \times 2n}, \quad E_n = \begin{bmatrix} & & 1 \\ & \ddots & \\ 1 & & \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

PROPOSITION 7.2. *Suppose  $\Pi \in \mathbb{C}^{n \times n}$  is unitary and satisfies  $\Pi^* = \epsilon_0 \Pi$  with  $\epsilon_0 = -1$  or  $\epsilon_0 = 1$ . Let  $\text{struct} = \text{struct}(\Pi, \star, \epsilon)$ . Then, for any unitarily invariant norm,*

$$K_{\text{struct}}(x, y) = K_{\Delta}(x, \Pi y),$$

where

$$(7.5) \quad \Delta = \begin{cases} \text{Herm} & \text{if } \epsilon_0 \epsilon = 1, \star = *, \\ \text{Sym}_{\mathbb{C}} & \text{if } \epsilon_0 \epsilon = 1, \star = \top, \\ \text{Sym}_{\mathbb{R}} & \text{if } \epsilon_0 \epsilon = 1, \star = \top, \text{ and } \Pi \in \mathbb{R}^{n \times n}, \\ \text{Skew}_{\mathbb{C}} & \text{if } \epsilon_0 \epsilon = -1, \star = \top, \\ \text{Skew}_{\mathbb{R}} & \text{if } \epsilon_0 \epsilon = -1, \star = \top, \text{ and } \Pi \in \mathbb{R}^{n \times n}. \end{cases}$$

Furthermore,  $K_{\text{struct}}(x, y) = K_{\text{Herm}}(x, i \Pi y)$  if  $\epsilon_0 \epsilon = -1$  and  $\star = *$ .

*Proof.* Using Lemma 7.1 and  $\Pi^*\Pi = I_n$  we obtain for the sets in (7.5)

$$\begin{aligned} K_{struct}(x, y) &= \{y^*\Delta x; \Delta \in struct, \|\Delta\| \leq 1\} \\ &= \{(\Pi y)^*(\Pi\Delta)x; \Pi\Delta \in \mathbf{\Delta}, \|\Pi\Delta\| \leq 1\} \\ &= K_{\mathbf{\Delta}}(x, \Pi y). \end{aligned}$$

The proof of the remaining statement is analogous.  $\square$

*Example 7.3.* In this final example we consider the matrix

$$(7.6) \quad A = \begin{bmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}, \quad \alpha, \beta \in \mathbb{R}, \beta \neq 0.$$

Note that  $A$  has two Jordan blocks of dimension 2 associated with the eigenvalues  $\lambda = \alpha + i\beta$ ,  $\bar{\lambda} = \alpha - i\beta$ . Both eigenvalues are nonderogatory and have algebraic multiplicity 2. Let  $x, \hat{y} \in \mathbb{C}^4$  be defined as

$$x := \frac{1}{\sqrt{2}} [1, i, 0, 0]^\top, \quad \hat{y} := \frac{1}{\sqrt{2}} [1, -i, 1, -i]^*.$$

Then

$$Ax = \lambda x, \quad \hat{y}^*(A - \lambda I_4)^2 = 0, \quad \hat{y}^*(A - \lambda I_4) \neq 0, \quad \hat{y}^*x = 1.$$

Let  $y^* := \hat{y}^*(A - \lambda I_4) = [0, 0, 1, -i]/\sqrt{2}$ . According to the discussion in the introduction we have for the connected component  $\mathcal{C}_{\mathbf{\Delta}}(A, \lambda, \epsilon)$  of the pseudospectrum  $\sigma_{\mathbf{\Delta}}(A, \epsilon)$  that

$$\lim_{\epsilon \rightarrow 0} \frac{\mathcal{C}_{\mathbf{\Delta}}(A, \lambda, \epsilon) - \lambda}{\epsilon^{1/2}} = K_{\mathbf{\Delta}}^{(2)}(x, y),$$

where  $K_{\mathbf{\Delta}}^{(2)}(x, y) = \{z; z^2 \in K_{\mathbf{\Delta}}(x, y)\}$ . We consider now the case where  $\mathbf{\Delta} = \mathbb{R}^{4 \times 4}$  and where the underlying norm is the Frobenius norm. Then by Theorem 6.5 the set  $K_{\mathbf{\Delta}}(x, y) = K_{\mathbb{R}^{4 \times 4}}(x, y)$  equals the ellipse defined in (6.1), where

$$a = \frac{1}{2} \|x\|^2 \|y\|^2 = \frac{1}{2}, \quad b = \frac{1}{2} (\overline{x^\top x}) (y^\top y) = 0.$$

Since  $b = 0$ , the ellipse is a disk of radius  $\sqrt{a} = 1/\sqrt{2}$ . Hence,  $K_{\mathbf{\Delta}}^{(2)}(x, y)$  is a disk of radius  $2^{-1/4}$ .

Note that the matrix  $A$  is Hamiltonian; i.e.,  $A$  is skew-adjoint with respect to the skew-Hermitian inner product  $(v, w) \mapsto v^* J w$ ,  $J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}$ . Let  $\text{Ham}_{\mathbb{R}} = \{\Delta \in \mathbb{R}^{4 \times 4}; \Delta^\top J = -J\Delta\}$  denote the set of all real Hamiltonian  $4 \times 4$  matrices. Then, by Lemma 7.1,  $\text{Ham}_{\mathbb{R}} = \{\Delta \in \mathbb{C}^{4 \times 4}; J\Delta \in \text{Sym}_{\mathbb{R}}\}$  and, by Proposition 7.2,  $K_{\text{Ham}_{\mathbb{R}}}(x, y) = K_{\text{Sym}_{\mathbb{R}}}(x, Jy)$ . Hence, if the underlying norm is the Frobenius norm, then  $K_{\text{Ham}_{\mathbb{R}}}(x, y)$  equals the ellipse with the parameters (see Theorem 6.5)

$$\begin{aligned} a &= \frac{1}{4} (\|x\|^2 \|Jy\|^2 + |x^\top Jy|^2) = \frac{1}{4}, \\ b &= \frac{1}{4} ((\overline{x^\top x}) ((Jy)^\top (Jy)) + (x^*(Jy))^2) = \frac{1}{4}. \end{aligned}$$

Since  $a = b$ , the ellipse degenerates to an interval. Precisely we have  $K_{\text{Ham}\mathbb{R}}(x, y) = [-\sqrt{a+b}, \sqrt{a+b}] = [-2^{-1/2}, 2^{-1/2}]$  and hence  $K_{\text{Ham}\mathbb{R}}^{(2)}(x, y) = [-2^{-1/4}, 2^{-1/4}] \cup i[-2^{-1/4}, 2^{-1/4}]$ .

**8. Appendix.** We give the proof Lemma 6.7. To this end we need the following fact.

**PROPOSITION 8.1.** *Let  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$  denote the singular values of  $M = AB^*$ , where  $A, B \in \mathbb{C}^{n \times r}$ . Then  $\sigma_k = 0$  for  $k > r$ , and  $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$  are the eigenvalues of  $(A^*A)(B^*B)$ . In particular,*

$$\sum_{k=1}^r \sigma_k^2 = \text{tr}((A^*A)(B^*B)), \quad \prod_{k=1}^r \sigma_k^2 = \det((A^*A)(B^*B)).$$

*Proof.* Since  $\text{rank}(M) \leq r$ , we have  $\sigma_k = 0$  for  $k > r$ . The squares of the singular values of  $M$  are the eigenvalues of  $M^*M = XY$ , where  $X = B$ ,  $Y = (A^*A)B^*$ . As is well known  $XY$  and  $YX = (A^*A)(B^*B)$  have the same nonzero eigenvalues.  $\square$

Now, let  $\sigma_1, \sigma_2$  denote the largest singular values of the matrix

$$M = a_1 b_1^* + a_2 b_2^* = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix}^*, \quad a_1, a_2, b_1, b_2 \in \mathbb{C}^n.$$

Since  $\text{rank}(M) \leq 2$ , the other singular values of  $M$  are zero. Using Proposition 8.1 we obtain for the Frobenius norm and the Schatten-1-norm of  $M$

$$\begin{aligned} \|M\|_F^2 &= \sigma_1^2 + \sigma_2^2 \\ &= \text{tr} \left( \begin{bmatrix} \|a_1\|^2 & a_1^* a_2 \\ a_2^* a_1 & \|a_2\|^2 \end{bmatrix} \begin{bmatrix} \|b_1\|^2 & b_1^* b_2 \\ b_2^* b_1 & \|b_2\|^2 \end{bmatrix} \right) \\ &= \|a_1\|^2 \|b_1\|^2 + \|a_2\|^2 \|b_2\|^2 + 2 \Re(\overline{a_1^* a_2} (b_1^* b_2)), \\ \|M\|_{(1)}^2 &= (\sigma_1 + \sigma_2)^2 \\ &= \sigma_1^2 + \sigma_2^2 + 2\sqrt{\sigma_1^2 \sigma_2^2} \\ &= \|M\|_F^2 + 2\sqrt{\beta}, \end{aligned}$$

where

$$\begin{aligned} \beta &= \det \left( \begin{bmatrix} \|a_1\|^2 & a_1^* a_2 \\ a_2^* a_1 & \|a_2\|^2 \end{bmatrix} \begin{bmatrix} \|b_1\|^2 & b_1^* b_2 \\ b_2^* b_1 & \|b_2\|^2 \end{bmatrix} \right) \\ &= (\|a_1\|^2 \|a_2\|^2 - |a_1^* a_2|^2)(\|b_1\|^2 \|b_2\|^2 - |b_1^* b_2|^2). \end{aligned}$$

Next, we compute the norms of the symmetric and the skew-symmetric parts of  $M$ . Let  $S_{\pm} = \frac{1}{2}(M \pm M^T)$ . Then  $S_{\pm}$  can be written in the form  $S_{\pm} = AB_{\pm}^*$ , where

$$A = \begin{bmatrix} a_1 & a_2 & \bar{b}_1 & \bar{b}_2 \end{bmatrix}, \quad B_{\pm} = \frac{1}{2} \begin{bmatrix} b_1 & b_2 & \pm \bar{a}_1 & \pm \bar{a}_2 \end{bmatrix}.$$

We have

$$A^*A = \begin{bmatrix} \|a_1\|^2 & a_1^*a_2 & \overline{a_1^\top b_1} & \overline{a_1^\top b_2} \\ a_2^*a_1 & \|a_2\|^2 & \overline{a_2^\top b_1} & \overline{a_2^\top b_2} \\ b_1^\top a_1 & b_1^\top a_2 & \|b_1\|^2 & \overline{b_1^\top b_2} \\ b_2^\top a_1 & b_2^\top a_2 & \overline{b_2^\top b_1} & \|b_2\|^2 \end{bmatrix},$$

$$B_\pm^*B_\pm = \frac{1}{4} \begin{bmatrix} \|b_1\|^2 & b_1^*b_2 & \pm \overline{b_1^\top a_1} & \pm \overline{b_1^\top a_2} \\ b_2^*b_1 & \|b_2\|^2 & \pm \overline{b_2^\top a_1} & \pm \overline{b_2^\top a_2} \\ \pm a_1^\top b_1 & \pm a_1^\top b_2 & \|a_1\|^2 & \overline{a_1^*a_2} \\ \pm a_2^\top b_1 & \pm a_2^\top b_2 & \overline{a_2^*a_1} & \|a_2\|^2 \end{bmatrix}.$$

Using Proposition 8.1 we obtain for the Frobenius norm of  $S_\pm$

$$\begin{aligned} \|S_\pm\|_F^2 &= \text{tr}((A^*A)(B_\pm^*B_\pm)) \\ &= \frac{1}{2} \left( \|a_1\|^2\|b_1\|^2 + \|a_2\|^2\|b_2\|^2 \pm |a_1^\top b_1|^2 \pm |a_2^\top b_2|^2 \right) \\ &\quad + \Re \left( (\overline{a_1^*a_2})(b_1^*b_2) \pm (\overline{a_1^\top b_2})(a_2^\top b_1) \right). \end{aligned}$$

We now determine the Schatten-1-norm of  $S_-$ . Since  $\text{rank}(S_-) \leq 4$ , at most four singular values of  $S_-$  are nonzero. Let  $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4$  denote these singular values. Since  $S_-$  is skew-symmetric, its singular values have even multiplicity [10, section 4.4, Exercise 26]. Thus  $\sigma_1 = \sigma_2$  and  $\sigma_3 = \sigma_4$ . This yields

$$\begin{aligned} \|S_-\|_{(1)}^2 &= (\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4)^2 \\ &= (2\sigma_1 + 2\sigma_3)^2 \\ &= 2(2\sigma_1^2 + 2\sigma_3^2) + 8\sigma_1\sigma_3 \\ &= 2\|S_-\|_F^2 + 8(\sigma_1^2\sigma_2^2\sigma_3^2\sigma_4^2)^{1/4} \\ &= 2\|S_-\|_F^2 + 8(\det[(A^*A)(B_-^*B_-)])^{1/4}. \end{aligned}$$

Since  $4\overline{B_-^*B_-}$  is unitarily similar to  $A^*A$ , we have  $\det(B_-^*B_-) = \frac{1}{256}\det(A^*A)$ . Hence,

$$\|S_-\|_{(1)}^2 = 2 \left( \|S_-\|_F^2 + \sqrt{\det(A^*A)} \right).$$

**Acknowledgment.** The author thanks the referees for many valuable comments.

#### REFERENCES

- [1] B. BERNHARDSSON, A. RANTZER, AND L. QIU, *Real perturbation values and real quadratic forms in a complex vector space*, Linear Algebra Appl., 270 (1998), pp. 131–154.
- [2] R. BHATIA, *Matrix Analysis*, Springer, New York, 1997.
- [3] S. BORA AND V. MEHRMANN, *Linear perturbation theory for structured matrix pencils arising in control theory*, SIAM J. Matrix Anal., 28 (2006), pp. 148–169.
- [4] R. BYERS AND D. KRESSNER, *On the condition of a complex eigenvalue under real perturbations*, BIT, 44 (2004), pp. 209–214.
- [5] F. CHAITIN-CHATELIN, A. HARRABI, AND A. ILAHI, *About Hölder condition numbers and the stratification diagram for defective eigenvalues*, Math. Comput. Simulation, 54 (2000), pp. 397–402.

- [6] M. EMBREE AND L. N. TREFETHEN, *The Pseudospectra Gateway*, <http://www.comlab.ox.ac.uk/pseudospectra>.
- [7] S. GRAILLAT AND F. TISSEUR, *Structured condition numbers and backward errors in scalar product spaces*, Electron. J. Linear Algebra, 15 (2006), pp. 159–177.
- [8] D. HINRICHSSEN AND A. J. PRITCHARD, *Mathematical Systems Theory I*, Texts Appl. Math. 48, Springer, Berlin, 2005.
- [9] J. HIRIAT-URRUTY AND C. LEMARECHAL, *Fundamentals of Convex Analysis*, Springer, Berlin, 2001.
- [10] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [11] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [12] M. KAROW, *Geometry of Spectral Value Sets*, Ph.D. thesis, University of Bremen, Bremen, Germany, 2003.
- [13] M. KAROW, D. KRESSNER, AND F. TISSEUR, *Structured eigenvalue condition numbers*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1052–1068.
- [14] T. KATO, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1980.
- [15] D. KRESSNER, M. J. PELÁEZ, AND J. MORO, *Structured Hölder condition numbers for multiple eigenvalues*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 175–201.
- [16] J. MORO, J. V. BURKE, AND M. L. OVERTON, *On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.
- [17] S. NOSCHESSE AND L. PASQUINI, *Eigenvalue condition numbers: Zero-structured versus traditional*, J. Comput. Appl. Math., 185 (2006), pp. 174–189.
- [18] S. NOSCHESSE AND L. PASQUINI, *Eigenvalue patterned condition numbers: Toeplitz and Hankel cases*, J. Comput. Appl. Math., 206 (2007), pp. 615–624.
- [19] J. R. RICE, *A theory of condition*, SIAM J. Numer. Anal., 3 (1966), pp. 287–310.
- [20] S. M. RUMP, *Eigenvalues, pseudospectrum and structured perturbations*, Linear Algebra Appl., 413 (2006), pp. 567–593.
- [21] S. M. RUMP AND H. SEKIGAWA, *The ratio between the Toeplitz and the unstructured condition number*, in Numerical Methods for Structured Matrices and Applications, Oper. Theory Adv. Appl. 199, Birkhäuser, Basel, 2010, pp. 397–419.
- [22] G. W. STEWART AND J. SUN, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.
- [23] F. TISSEUR, *A chart of backward errors for singly and doubly structured eigenvalue problems*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 877–897.
- [24] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.